

PATENT**ULTRA LOW LEAKAGE MOSFET TRANSISTOR**

Inventors: Peter J. Hopper, San Jose, California
Philip Lindorfer, San Jose, California
5 Vladislav Vashchenko, Palo Alto, California

TECHNICAL FIELD

The present invention relates to a design for an ultra low leakage MOSFET transistor that is formed by minimizing band-to-band and trap assisted tunneling mechanisms leading to gate induced drain leakage characteristics.

BACKGROUND OF THE INVENTION

MOSFET transistors are widely used in semiconductor circuits. Figs. 1A and 1B show a conventional MOSFET transistor structure. Fig. 1A shows a top view of the MOSFET transistor 100; Fig. 1B shows a side view taken along line A-A'. Spaced-apart n-type source region 110 and n-type drain region 112 formed in p-type semiconductor substrate 114 define a substrate channel region 116. A drain contact 106 and a source contact 108 are also provided. A conductive gate electrode 104 is disposed over the channel 116, with a gate oxide layer 120 disposed between the gate 104 and the channel 116.

Depending upon a voltage applied to the gate 104, the channel 116 can be made to become conductive or nonconductive. When the channel is conductive, a voltage applied across the drain contact 106 and the source contact 108 results in electrons flowing from the source 110 to the drain 112. The amount of current flowing between the source 110 and drain 112 is proportional to the amount of resistance of the channel region 116. As shown in Figs. 1A and 1B, the structure of a conventional MOSFET is such that the length (Lc1) of the gate 104 is constant across the area separating the source 110 and drain 112.

Since the process used to fabricate a MOSFET structure is well known, the entire process is not described in detail herein; relevant aspects of the process are provided. For an N-channel MOSFET, a P type silicon substrate 114 is used. The region of the substrate 114 in which the MOSFET is formed can be isolated using shallow trench isolation techniques (STI), or other isolation processes such a local oxidation of silicon (LOCOS) process. In the

STI process, for example, a photoresist mask can be formed over the substrate 114 and trench is etched into the surface of the substrate 114. The trench is then filled with an insulating material such as SiO₂. The resulting isolation region 102 is shown in Figs. 1A and 1B. The isolation region 102 operates to isolate the MOSFET structure from other devices formed in an integrated circuit. The isolation region 102 forms a sidewall along the source 110, drain 112 and on both ends of the channel 104. This sidewall area has discontinuities, or traps, in the structure of the silicon substrate 114, which are due to stresses on the substrate resulting from the process used to form the sidewall of the isolation area 102.

As shown in Fig. 1B, an insulating layer 120 is disposed on top of the substrate 114. The gate 104, which can be formed of polysilicon, or other suitable conductive material, is formed on top of the insulating layer 120. For NMOS devices, typically after the gate 104 has been formed, N type atoms, such as Arsenic, are implanted into the substrate 114 to form the source region 110 and drain region 112 and to dope the gate polysilicon. The area under the gate 104 then forms the channel 116. Thus, the channel region 116 provides a uniform distance (Lc1) separating the source 110 from the drain 112.

As shown in Fig. 1A, current 118 flowing between the drain 112 and source 110 results when a voltage differential is applied across the drain contact 106 and the source contact 108 and an appropriate voltage has also been applied to the gate 104. Given that the length Lc1 of the channel 116 separating the source 110 from the drain 112 is constant across the width of the device, the resistance across the width of the device is uniform.

Electrons moving through the channel 116 between the source 110 and drain 112 which are in close proximity to the sidewall with the isolation region 102 will tend to have their movement impeded by the traps that result from the formation of the trench. This results in the generation of low frequency noise, some times referred to as 1/f noise. The channel region 116 includes two principal 1/f noise sources: those originating from the stress induced along the isolation sidewall and those originating from the silicon dioxide to silicon interface under the poly. The larger the amount of current flowing along edge of the channel region 116 that is adjacent to the sidewall of the isolation region 102, the greater the noise generated.

Co-pending and commonly-assigned U.S. Patent Application Serial No. 10/423,065, filed on April 25, 2003, discloses the MOSFET structure 200 shown in Figs. 2A-E. As shown therein, a drain region 212 is provided having a drain contact 206. The MOSFET 200 also includes a source region 210 and having a source contact 208. A channel region 216 is disposed between the source 210 and drain 212. A gate 204 is disposed above the channel region 216 and is separated from the channel 216 by a gate oxide layer 220. However, unlike in the MOSFET structure discussed above with respect to Figs. 1A and 1B, the length of the gate 204 is not the same across the width of the MOSFET 200.

As shown in Fig. 2A, the area of the gate 204 shown as having a width of W_2 corresponds to the area over the isolation area sidewall 201 and over the area of the channel 216 that is adjacent to the sidewall 201. This area has a width of W_2 and a length of L_2 . As further shown in Fig. 2A, a second area of the gate 204 is above a region where the channel 216 is adjacent to the sidewall 205 of the isolation area 202. In this area above the region of the sidewall 205, the gate 204 has a width shown as W_3 and a length of L_2 . In between the areas of the gate 204 which are above the sidewalls 201 and 205, another area of the gate is provided which has a width of W_1 and a length of L_1 .

The process used to form the MOSFET 200 is very similar to that described above in connection with conventional MOSFETs. The difference is in the shape of the gate 204 of the MOSFET 200. Further, given that the shape of gate 204 is different, the channel region 216 is also of a different shape than the prior MOSFET channel regions. The process used to form the drain 212 and source 210 regions is a self-aligned process wherein atoms are implanted into the substrate 214 to form the source 210 and drain 212. The gate 204 is formed above the substrate 214 prior to implanting the atoms that form the source 210 and drain 212. Thus, when the atoms are implanted to form the drain 212 and the source 210, the material of the gate 204, which can be polysilicon or other conductive material, operates to prevent the atoms from entering the substrate 214 region underlying the gate, and this region forms the channel 216.

As shown in Fig. 2A, the area of the gate 204 having a width shown as W_2 has a length of L_2 , which is substantially greater than the length of the gate in the area of the gate 204 where the width of the gate is W_1 and the length is L_1 . Correspondingly, the channel region 216 under the area of the gate 204 having a width of W_2 has a length L_2 , which is

much longer than the length $L1$ in the area where the width of the gate 204 is $W1$. Further, the area of the gate 204 where the width is $W3$ has a length of $L2$.

Fig. 2B shows a cross section of the MOSFET 200 taken along cutaway lines 2B-2B' in Fig. 2A. This cross section view shows the isolation areas 202 disposed on either side of the MOSFET drain 212, source 210 and channel 216. Specifically, the sidewall 201 of the isolation area 202 is adjacent to one end of the channel 216, where the end of the channel 216 adjacent to the sidewall 201 has a width of $W1$; and the sidewall 205 of the isolation area 202 is adjacent to the other end of the channel 216, where the end of the channel adjacent 216 to the sidewall 205 has a width of $W3$. As shown along cross section 2B-2B', an area of the source 210 is between areas of the channel 216. Further, the gate 204 is disposed over the channel 216. To further illustrate the structure of the MOSFET, Fig. 2C shows a cross sectional view of the MOSFET 200 along cutaway lines 2C-2C' in Fig. 2B. As shown along the center of the MOSFET 200, the channel 216 is continuous between sidewall 201 of the isolation area 202 and sidewall 205. The gate 204 is disposed over the channel 216.

Fig. 2D shows a cross sectional view of the MOSFET 200 along cutaway lines 2D-2D' in Fig. 2A. As shown in Fig. 2D, the drain 212 is separated from the source 210 by the channel 216. The channel 216 in the area shown has a length of $L2$. The width of the channel 216 and the gate 204 in this area is $W3$.

Fig. 2E shows a cross section view of the MOSFET 200 along cutaway lines 2E-2E' in Fig. 2A. As shown in Fig. 2E, the drain 212 is separated from the source 210 by the channel 216. The channel along the line 2E-2E' has a length of $L1$.

The operation of the MOSFET 200 is similar to that discussed above with respect to the prior MOSFET 100. However, the different shape of the gate 204 and channel 216 in the MOSFET 200 result in beneficial characteristics. To consider the operation of the MOSFET 200, it is helpful to recognize the MOSFET 200 as having 3 parts. One part corresponds to the drain 212, the channel 216 and the source 210 where the channel has a width of $W2$ and a length of the $L2$. In this area, in response to voltage being applied to the gate 204, the channel 216 will transport electrons when a voltage differential is applied across the drain contact 206 and the source contact 208. The amount of current flowing through the channel 216 in this area along the sidewall 201 of the isolation area 202 is proportional to

W2/L2. The current flowing through center area of the channel 216 having a width W1 is proportional to $W1/L1$. The current flowing through the channel 216 in the area along the sidewall 205 of the isolation area 202 is proportional to $W3/L2$. By providing a gate 204, and underlying channel 216 shaped to increase the amount of current flowing through a center part of the channel 214 and to decrease the amount of current flowing in an area along the sidewalls 201 and 205 of the isolation area 202, the amount of noise generated by current flowing through the MOSFET 200 is minimized. As discussed above, the stresses to the underlying substrate 214 induced by the forming of the sidewall 201 and 205 of the isolation area 202 create discontinuities or traps which interfere with the movement of electrons through the channel 216 in the areas proximate to the sidewalls 201 and 205. This results in noise (1/f noise). The amount of noise is proportional to the amount of current flowing along the sidewall.

MOSFET leakage is derived from a series of physical mechanisms. To design a low leakage transistor, each of these mechanisms should be minimized. Typically, and primarily, CMOS device leakage mechanisms are dominated by drain induced barrier lowering (DIBL) and by gate induced drain leakage (GIDL). DIBL may be simply minimized by minimizing short channel effects, enlarging channel lengths and raising channel dopant concentrations. Once DIBL is minimized to below a picoamp per micro width, leakage is typically governed by the level of GIDL.

GIDL is associated with a tunneling mechanism and, in turn, is dominated by trap assisted tunneling. Thus, if the level of traps associated with the semiconductor region under the direct influence of a gate-drain Efield, then the level of GIDL leakage will be lowered and, in turn, provide the path to a low leakage transistor.

25 BRIEF DESCRIPTION OF THE DRAWINGS

Figs. 1A-1B show views of a conventional MOSFET transistor structure.

Figs. 2 A-2E show views of an alternate MOSFET transistor structure.

Figs. 3A-3G show views of an embodiment of a MOSFET transistor structure in accordance with the present invention.

30 Fig. 4A shows a layout of a CMOS transistor structure that utilizes the concepts of the present invention.

Fig. 4B is a schematic diagram illustrating a CMOS circuit corresponding to the Fig. 4A structure.

DETAILED DESCRIPTION

5 Figs. 3A and 3B show a MOSFET transistor structure 300 in accordance with the concepts of the present invention. Although the transistor structure 300 is described in the context of an NMOS device, those skilled in the art will appreciate that the concepts of the invention are equally applicable to PMOS devices.

The transistor structure 300 is formed in a substrate of semiconductor material 302,
10 preferably silicon, having a first conductivity type. Perimeter isolation dielectric material 304, which preferably is shallow trench isolation (STI) silicon dioxide, but can be LOCOS (local oxidation of silicon) or other types of conventional isolation structures, is formed in the substrate 302 along the entire perimeter of an active region 306 to define a continuous sidewall interface 308 between the isolation dielectric material 304 and the active region 306.
15 Spaced-apart source and drain regions 310, 312 having a second conductivity type that is opposite the first conductivity type are formed in the active region 306 to define a substrate channel region 314 therebetween. (As discussed in greater detail below, both the source region 310 and the drain region 312 preferably include a low density diffusion (LDD) region and deep implant region). As further shown in Figs. 3A and 3B, both the source region 310
20 and the drain region 312 are also spaced-apart from the isolation dielectric material 304. A conductive gate electrode 316 that includes a first portion 316a that extends over the substrate channel region 314 and a second portion 316b that extends continuously over the entire sidewall interface 308 between the isolation dielectric material 304 and the active device region 306 is separated from the active region 306 by intervening gate dielectric
25 material 318.

The MOSFET transistor structure shown in Figs. 3A and 3B can be formed in accordance with the following integrated circuit manufacturing process steps illustrated in Figs. 3C-3G.

First, as shown in Figs. 3C and 3D, isolation dielectric material 304, shown in this
30 example as STI, is formed in the semiconductor substrate 302 (e.g. silicon) such that the isolation dielectric material 304 defines an active region 306 of the substrate 302. The

isolation material 304 is formed along the entire perimeter of the active region 306 to define a continuous sidewall interface 308 between the isolation dielectric material 304 and the active region 306.

As shown in Fig. 3E, a layer of gate dielectric material 318, e.g. silicon dioxide, is then formed to extend over the active region 306 and over the continuous sidewall interface 308 between the isolation dielectric material 304 and the active region 302. A layer of conductive material 316, e.g. polysilicon, is then formed over the gate dielectric material 318.

As shown in Fig. 3F, a mask 320, typically photoresist, is then formed over the conductive gate material 316 and patterned to define the source and drain openings. The mask 320 is then utilized to etch the layer of conductive gate material 316 to provide a first "channel" portion 316a of the conductive gate 316 and a second "perimeter" portion 316b of the conductive gate 316. As further shown in Fig. 3F, this etch process can result in etching of exposed regions of the gate dielectric material 318.

Referring to Fig. 3G, following the patterning of the conductive gate material 316 to define the first portion 316a and the second portion 316b as discussed above, a low density diffusion (LDD) step is performed to introduce dopant material having a second conductivity type into the source region 310 to provide source LDD region 310a and into drain region 312 to provide drain LDD region 312a. Following the LDD step, a layer of silicon oxide is formed and then anisotropically etched to provide oxide spacers 322 on the sidewalls of the conductive gate 316. A second, deeper implant of dopant material of the second conductivity type is then performed to provide the deep source region 310b and the deep drain region 312b. Since the implant of the LDD regions 310a, 312a is self-aligned to the conductive gate 316, which extends over the sidewall interface 308 into the active region 306 along the entire perimeter of the active region 306, and the deep implant is self-aligned to the sidewall spacers 322, both the source region 310 and the drain region 312 are spaced-apart from the sidewall interface 308 between the isolation dielectric material 304 and the active region 306.

The MOSFET structure is then completed utilizing well-known processing steps.

Thus, a new form of MOSFET transistor structure is provided for use with either local or trench or other isolation schemes. Instead of the gate polysilicon being drawn across an active region, as is typical in CMOS devices as shown in Figs. 1A and 1B, or as has been

proposed as a low 1/f noise MOSFET of the type shown in Figs. 2A and 2B, an enclosure ring is maintained around the entire composite perimeter, thereby completely avoiding the regions of high trap density and, thus, preventing any current path for the GIDL mechanism to take place.

5 The premise for this proposal is based upon the fact that a higher density of traps is associated with the perimeter of an active region. This is due to the perimeter's association with stress induced crystal slippage, crystal strain, silicon/silicon dioxide interfacial traps, intrinsic gettering properties for glsile defects, oxidation induced precipitates and even dislocations.

10 Further, it should be noted that the GIDL mechanism is not one associated with source to drain leakage.

 Since process variations and differences in device architecture can result in differing trap densities along the sidewall interface between the isolation material and the device
15 active region, those skilled in the art will appreciate that the width of the enclosure ring provided between the isolation sidewall and the source and drain diffusion regions in a MOSFET structure in accordance with the present invention will vary depending upon the application. For those devices, e.g. imaging devices, that are particularly sensitive to the GIDL mechanism, it is believed that the width of the enclosure ring should be about 1000-
20 5000 Angstroms.

 Fig. 4A shows the application of the concepts of the present invention to a CMOS transistor structure 400 that includes a PMOS device and an NMOS device. Those skilled in the art will appreciate that the NMOS device is formed in a p-type silicon substrate, while the
25 PMOS device is formed in an n-well formed in the p-type substrate. As shown in Fig. 4A, isolation oxide (STI) formed in a silicon substrate defines the active device region of the CMOS structure 400. Spaced-apart p+ source and drain regions 402, 404 and intervening n+ channel region define the PMOS device. Spaced-apart n+ source and drain regions 406, 408 and intervening p+ channel define the NMOS device. Continuous polysilicon gate 410
30 provides the gate electrode for both the PMOS device and the NMOS device as well as a perimeter portion that overlaps the sidewall interface between the STI and the active device

region to ensure a self-aligned source/drain implant that maintains an enclosure ring around the entire composite to prevent a GIDL current path. (Fig. 4B provides a simple CMOS circuit schematic.)

- 5 Although only specific embodiments of the present invention are shown and described herein, the invention is not to be limited by these embodiments. Rather, the scope of the invention is to be defined by these descriptions taken together with the attached claims and their equivalents.